

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Deboer, Herman A.
Strijker, Rein
Heyneker, Herbert L.
Platenburg, Gerald
Lee, Sang He
Pieper, Frank
Krimpenfort, Paul J.A.

(ii) TITLE OF INVENTION: Method of Producing a Transgenic Bovine or Transgenic Bovine Embryo (as amended)

(iii) NUMBER OF SEQUENCES: 38

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Townsend and Townsend and Crew LLP
(B) STREET: Two Embarcadero Center, 8th Floor
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94111-3834

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patent In Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/154,019
(B) FILING DATE: 16-NOV-1993
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/077,788
(B) FILING DATE: 15-JUN-1993

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/898,956
(B) FILING DATE: 15-JUN-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/619,131
(B) FILING DATE: 27-NOV-1990

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/444,745
(B) FILING DATE: 01-DEC-1989

ac
3/28/97

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Liebescheutz, Joe O.
- (B) REGISTRATION NUMBER: 37,505
- (C) REFERENCE/DOCKET NUMBER: 16994-003122

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 415-326-2400
- (B) TELEFAX: 415-326-2422

du
3/28/97

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2319 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 1..54

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 55..2130

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2130

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGA CTT GTC TTC CTC GTC CTG CTG TTC CTC GGG GCC CTC GGA CTG TGT
 48
 Gly Leu Val Phe Leu Val Leu Leu Phe Leu Gly Ala Leu Gly Leu Cys
 -18 -15 -10 -5

CTG GCT GGC CGT AGG AGA AGG AGT GTT CAG TGG TGC GCC GTA TCC CAA
 96
 Leu Ala Gly Arg Arg Arg Arg Ser Val Gln Trp Cys Ala Val Ser Gln
 1 5 10

CCC GAG GCC ACA AAA TGC TTC CAA TGG CAA AGG AAT ATG AGA AAA GTG
 144
 Pro Glu Ala Thr Lys Cys Phe Gln Trp Gln Arg Asn Met Arg Lys Val
 15 20 25 30

CTG GGC CCT CCT GTC AGC TGC ATA AAG AGA GAC TCC CCC ATC CAG TGT
 192
 Leu Gly Pro Pro Val Ser Cys Ile Lys Arg Asp Ser Pro Ile Gln Cys
 35 40 45

ATC CAG GCC ATT GCG GAA AAC AGG GCC GAT GCT GTG ACC CTT GAT GGT
 240
 Ile Gln Ala Ile Ala Glu Asn Arg Ala Asp Ala Val Thr Leu Asp Gly

du
3/24/97

50

55

60

GGT TTC ATA TAC GAG GCA GGC CTG GCC CCC TAC AAA CTG CGA CCT GTA
288

Gly Phe Ile Tyr Glu Ala Gly Leu Ala Pro Tyr Lys Leu Arg Pro Val
65 70 75

GCG GCG GAA GTC TAC GGG ACC GAA AGA CAG CCA CGA ACT CAC TAT TAT
336

Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr Tyr
80 85 90

de
3/26/97

124

CGG GTG GCT GTG GTG AAG AAG GGC GGC AGC TTT CAG CTG AAC GAA CTG
384
Arg Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu Leu
95 100 105 110

CAA GGT CTG AAG TCC TGC CAC ACA GGC CTT CGC AGG ACC GCT GGA TGG
432
Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly Trp
115 120 125

AAT GTC CCT ACA GGG ACA CTT CGT CCA TTC TTG AAT TGG ACG GGT CCA
480
Asn Val Pro Thr Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro
130 135 140

CCT GAG CCC ATT GAG GCA GCT GTG CAG TTC TTC TCA GCC AGC TGT GTT
528
Pro Glu Pro Ile Glu Ala Ala Val Gln Phe Phe Ser Ala Ser Cys Val
145 150 155

CCC GGT GCA GAT AAA GGA CAG TTC CCC AAC CTG TGT CGC CTG TGT GCG
576
Pro Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys Ala
160 165 170

GGG ACA GGG GAA AAC AAA TGT GCC TTC TCC TCC CAG GAA CCG TAC TTC
624
Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr Phe
175 180 185 190

AGC TAC TCT GGT GCC TTC AAG TGT CTG AGA GAC GGG GCT GGA GAC GTG
672
Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp Val
195 200 205

GCT TTT ATC AGA GAG AGC ACA GTG TTT GAG GAC CTG TCA GAC GAG GCT
720
Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu Ala
210 215 220

GAA AGG GAC GAG TAT GAG TTA CTC TGC CCA GAC AAC ACT CGG AAG CCA
768
Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys Pro
225 230 235

125

dc
3/28/97

125

GTG GAC AAG TTC AAA GAC TGC CAT CTG GCC CGG GTC CCT TCT CAT GCC
816
Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His Ala
240 245 250

GTT GTG GCA CGA AGT GTG AAT GGC AAG GAG GAT GCC ATC TGG AAT CTT
864
Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn Leu
255 260 265 270

CTC CGC CAG GCA CAG GAA AAG TTT GGA AAG GAC AAG TCA CCG AAA TTC
912
Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys Phe
275 280 285

CAG CTC TTT GGC TCC CCT AGT GGG CAG AAA GAT CTG CTG TTC AAG GAC
960
Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys Asp
290 295 300

RL

3/28/97

126

TCT GCC ATT GGG TTT TCG AGG GTG CCC CCG AGG ATA GAT TCT GGG CTG
1008
Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly Leu
305 310 315

TAC CTT GGC TCC GGC TAC TTC ACT GCC ATC CAG AAC TTG AGG AAA AGT
1056
Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys Ser
320 325 330

GAG GAG GAA GTG GCT GCC CGG CGT GCG CGG GTC GTG TGG TGT GCG GTG
1104
Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala Val
335 340 345 350

GGC GAG CAG GAG CTG CGC AAG TGT AAC CAG TGG AGT GGC TTG AGC GAA
1152
Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser Glu
355 360 365

GGC AGC GTG ACC TGC TCC TCG GCC TCC ACC ACA GAG GAC TGC ATC GCC
1200
Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile Ala
370 375 380

CTG GTG CTG AAA GGA GAA GCT GAT GCC ATG AGT TTG GAT GGA GGA TAT
1248
Leu Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Tyr
385 390 395

GTG TAC ACT GCA TGC AAA TGT GGT TTG GTG CCT GTC CTG GCA GAG AAC
1296
Val Tyr Thr Ala Cys Lys Cys Gly Leu Val Pro Val Leu Ala Glu Asn
400 405 410

TAC AAA TCC CAA CAA AGC AGT GAC CCT GAT CCT AAC TGT GTG GAT AGA
1344
Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val Asp Arg
415 420 425 430

CCT GTG GAA GGA TAT CTT GCT GTG GCG GTG GTT AGG AGA TCA GAC ACT
1392
Pro Val Glu Gly Tyr Leu Ala Val Ala Val Val Arg Arg Ser Asp Thr
435 440 445

AGC CTT ACC TGG AAC TCT GTG AAA GGC AAG AAG TCC TGC CAC ACC GCC

127

127

20
3/28/97

1440
Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His Thr Ala
450 455 460

GTG GAC AGG ACT GCA GGC TGG AAT ATC CCC ATG GGC CTG CTC TTC AAC
1488

Val Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly Leu Leu Phe Asn
465 470 475

CAG ACG GGC TCC TGC AAA TTT GAT GAA TAT TTC AGT CAA AGC TGT GCC
1536

Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser Cys Ala
480 485 490

CCT GGG TCT GAC CCG AGA TCT AAT CTC TGT GCT CTG TGT ATT GGC GAC
1584

Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile Gly Asp
495 500 505 510

128

de
3/28/97

128

GAG CAG GGT GAG AAT AAG TGC GTG CCC AAC AGC AAT GAG AGA TAC TAC
1632
Glu Gln Gly Glu Asn Lys Cys Val Pro Asn Ser Asn Glu Arg Tyr Tyr
515 520 525

GGC TAC ACT GGG GCT TTC CGG TGC CTG GCT GAG AAT GCT GGA GAC GTT
1680
Gly Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asn Ala Gly Asp Val
530 535 540

GCA TTT GTG AAA GAT GTC ACT GTC TTG CAG AAC ACT GAT GGA AAT AAC
1728
Ala Phe Val Lys Asp Val Thr Val Leu Gln Asn Thr Asp Gly Asn Asn
545 550 555

AAT GAG GCA TGG GCT AAG GAT TTG AAG CTG GCA GAC TTT GCG CTG CTG
1776
Asn Glu Ala Trp Ala Lys Asp Leu Lys Leu Ala Asp Phe Ala Leu Leu
560 565 570

TGC CTC GAT GGC AAA CGG AAG CCT GTG ACT GAG GCT AGA AGC TGC CAT
1824
Cys Leu Asp Gly Lys Arg Lys Pro Val Thr Glu Ala Arg Ser Cys His
575 580 585 590

CTT GCC ATG GCC CCG AAT CAT GCC GTG GTG TCT CGG ATG GAT AAG GTG
1872
Leu Ala Met Ala Pro Asn His Ala Val Val Ser Arg Met Asp Lys Val
595 600 605

GAA CGC CTG AAA CAG GTG CTG CTC CAC CAA CAG GCT AAA TTT GGG AGA
1920
Glu Arg Leu Lys Gln Val Leu Leu His Gln Gln Ala Lys Phe Gly Arg
610 615 620

AAT GGA TCT GAC TGC CCG GAC AAG TTT TGC TTA TTC CAG TCT GAA ACC
1968
Asn Gly Ser Asp Cys Pro Asp Lys Phe Cys Leu Phe Gln Ser Glu Thr
625 630 635

AAA AAC CTT CTG TTC AAT GAC AAC ACT GAG TGT CTG GCC AGA CTC CAT
2016
Lys Asn Leu Leu Phe Asn Asp Asn Thr Glu Cys Leu Ala Arg Leu His
640 645 650

GGC AAA ACA ACA TAT GAA AAA TAT TTG GGA CCA CAG TAT GTC GCA GGC

129

1282

2064

Gly Lys Thr Thr Tyr Glu Lys Tyr Leu Gly Pro Gln Tyr Val Ala Gly
655 660 665 670

ATT ACT AAT CGT AAA AAG TGC TCA ACC TCC CCC CTC CTG GAA GCC TGT
2112

Ile Thr Asn Arg Lys Lys Cys Ser Thr Ser Pro Leu Leu Glu Ala Cys
675 680 685

GAA TTC CTC AGG AAG TAAAACCGAA GAAGATGGCC CAGCTCCCCA AGAAAGCCTC
2167

Glu Phe Leu Arg Lys
690

AGCCATTCAC TGCCCCCAGC TCTTCTCCCC AGGTGTGTTG GGGCCTTGGC TCCCCTGCTG
2227

AAGGTGGGGA TTGCCCATCC ATCTGCTTAC AATTCCTGC TGTCGTCTTA GCAAGAAGTA
2287

AAATGAGAAA TTTTGTTGAT ATTCAAAAAA AA
2319

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 709 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly Leu Val Phe Leu Val Leu Leu Phe Leu Gly Ala Leu Gly Leu Cys
 -18 -15 -10 -5

Leu Ala Gly Arg Arg Arg Arg Ser Val Gln Trp Cys Ala Val Ser Gln
 1 5 10

Pro Glu Ala Thr Lys Cys Phe Gln Trp Gln Arg Asn Met Arg Lys Val
 15 20 25 30

Leu Gly Pro Pro Val Ser Cys Ile Lys Arg Asp Ser Pro Ile Gln Cys
 35 40 45

Ile Gln Ala Ile Ala Glu Asn Arg Ala Asp Ala Val Thr Leu Asp Gly
 50 55 60

Gly Phe Ile Tyr Glu Ala Gly Leu Ala Pro Tyr Lys Leu Arg Pro Val
 65 70 75

Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr Tyr
 80 85 90

Arg Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu Leu
 95 100 105 110

Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly Trp
 115 120 125

Asn Val Pro Thr Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro
 130 135 140

Pro Glu Pro Ile Glu Ala Ala Val Gln Phe Phe Ser Ala Ser Cys Val
 145 150 155

3/28/97

130

Pro Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys Ala
160 165 170

Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr Phe
175 180 185 190

Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp Val
195 200 205

132

de
3/28/98

13/

Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu Ala
210 215 220

Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys Pro
225 230 235

Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His Ala
240 245 250

Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn Leu
255 260 265 270

Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys Phe
275 280 285

Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys Asp
290 295 300

Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly Leu
305 310 315

Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys Ser
320 325 330

Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala Val
335 340 345 350

Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser Glu
355 360 365

Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile Ala
370 375 380

Leu Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Tyr
385 390 395

Val Tyr Thr Ala Cys Lys Cys Gly Leu Val Pro Val Leu Ala Glu Asn
400 405 410

Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val Asp Arg
415 420 425 430

dc
3/28/97

132

Pro Val Glu Gly Tyr Leu Ala Val Ala Val Val Arg Arg Ser Asp Thr
435 440 445

Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His Thr Ala
450 455 460

Val Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly Leu Leu Phe Asn
465 470 475

Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser Cys Ala
480 485 490

134

de 3/28/97

133

Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile Gly Asp
495 500 505 510

Glu Gln Gly Glu Asn Lys Cys Val Pro Asn Ser Asn Glu Arg Tyr Tyr
515 520 525

Gly Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asn Ala Gly Asp Val
530 535 540

Ala Phe Val Lys Asp Val Thr Val Leu Gln Asn Thr Asp Gly Asn Asn
545 550 555

Asn Glu Ala Trp Ala Lys Asp Leu Lys Leu Ala Asp Phe Ala Leu Leu
560 565 570

Cys Leu Asp Gly Lys Arg Lys Pro Val Thr Glu Ala Arg Ser Cys His
575 580 585 590

Leu Ala Met Ala Pro Asn His Ala Val Val Ser Arg Met Asp Lys Val
595 600 605

Glu Arg Leu Lys Gln Val Leu Leu His Gln Gln Ala Lys Phe Gly Arg
610 615 620

Asn Gly Ser Asp Cys Pro Asp Lys Phe Cys Leu Phe Gln Ser Glu Thr
625 630 635

Lys Asn Leu Leu Phe Asn Asp Asn Thr Glu Cys Leu Ala Arg Leu His
640 645 650

Gly Lys Thr Thr Tyr Glu Lys Tyr Leu Gly Pro Gln Tyr Val Ala Gly
655 660 665 670

Ile Thr Asn Arg Lys Lys Cys Ser Thr Ser Pro Leu Leu Glu Ala Cys
675 680 685

Glu Phe Leu Arg Lys
690

135

dc
3/28/97

134

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2619 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 295..351

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 352..2430

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 295..2430

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACTCCTAGG GGCTTGCAGA CCTAGTGGGA GAGAAAGAAC ATCGCAGCAG CCAGGCAGAA
60

CCAGGACAGG TGAGGTGCAG GCTGGCTTTC CTCTCGCAGC GCGGTGTGGA GTCCTGTCCT
120

GCCTCAGGGC TTTTCGGAGC CTGGATCCTC AAGGAACAAG TAGACCTGGC CGCGGGGAGT
180

GGGGAGGGAA GGGGTGTCTA TTGGGCAACA GGGCGGCAAA GCCCTGAATA AAGGGGCGCA
240

GGGCAGGCGC AAGTGCAGAG CCTTCGTTTG CCAAGTCGCC TCCAGACCGC AGAC ATG
297

Met
-19

AAA CTT GTC TTC CTC GTC CTG CTG TTC CTC GGG GCC CTC GGA CTG TGT
345

Lys Leu Val Phe Leu Val Leu Leu Phe Leu Gly Ala Leu Gly Leu Cys
-15 -10 -5

CTG GCT GGC CGT AGG AGA AGG AGT GTT CAG TGG TGC GCC GTA TCC CAA
393

Leu Ala Gly Arg Arg Arg Arg Ser Val Gln Trp Cys Ala Val Ser Gln
1 5 10

136

3/28/97

135

CCC GAG GCC ACA AAA TGC TTC CAA TGG CAA AGG AAT ATG AGA AAA GTG
441
Pro Glu Ala Thr Lys Cys Phe Gln Trp Gln Arg Asn Met Arg Lys Val
15 20 25 30

CGT GGC CCT CCT GTC AGC TGC ATA AAG AGA GAC TCC CCC ATC CAG TGT
489
Arg Gly Pro Pro Val Ser Cys Ile Lys Arg Asp Ser Pro Ile Gln Cys
35 40 45

3/28/97

136

ATC CAG GCC ATT GCG GAA AAC AGG GCC GAT GCT GTG ACC CTT GAT GGT
 537
 Ile Gln Ala Ile Ala Glu Asn Arg Ala Asp Ala Val Thr Leu Asp Gly
 50 55 60

GGT TTC ATA TAC GAG GCA GGC CTG GCC CCC TAC AAA CTG CGA CCT GTA
 585
 Gly Phe Ile Tyr Glu Ala Gly Leu Ala Pro Tyr Lys Leu Arg Pro Val
 65 70 75

GCG GCG GAA GTC TAC GGG ACC GAA AGA CAG CCA CGA ACT CAC TAT TAT
 633
 Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr Tyr
 80 85 90

GCC GTG GCT GTG GTG AAG AAG GGC GGC AGC TTT CAG CTG AAC GAA CTG
 681
 Ala Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu Leu
 95 100 105 110

CAA GGT CTG AAG TCC TGC CAC ACA GGC CTT CGC AGG ACC GCT GGA TGG
 729
 Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly Trp
 115 120 125

AAT GTC CCT ACA GGG ACA CTT CGT CCA TTC TTG AAT TGG ACG GGT CCA
 777
 Asn Val Pro Thr Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro
 130 135 140

CCT GAG CCC ATT GAG GCA GCT GTG GCC AGG TTC TTC TCA GCC AGC TGT
 825
 Pro Glu Pro Ile Glu Ala Ala Val Ala Arg Phe Phe Ser Ala Ser Cys
 145 150 155

GTT CCC GGT GCA GAT AAA GGA CAG TTC CCC AAC CTG TGT CGC CTG TGT
 873
 Val Pro Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys
 160 165 170

GCG GGG ACA GGG GAA AAC AAA TGT GCC TTC TCC TCC CAG GAA CCG TAC
 921
 Ala Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr
 175 180 185 190

TTC AGC TAC TCT GGT GCC TTC AAG TGT CTG AGA GAC GGG GCT GGA GAC
 969
 Phe Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp
 195 200 205

GTG GCT TTT ATC AGA GAG AGC ACA GTG TTT GAG GAC CTG TCA GAC GAG
 1017
 Val Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu

3/28/97

137

210

215

220

GCT GAA AGG GAC GAG TAT GAG TTA CTC TGC CCA GAC AAC ACT CGG AAG
1065

Ala Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys
225 230 235

CCA GTG GAC AAG TTC AAA GAC TGC CAT CTG GCC CGG GTC CCT TCT CAT
1113

Pro Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His
240 245 250

3/28/97

138

GCC GTT GTG GCA CGA AGT GTG AAT GGC AAG GAG GAT GCC ATC TGG AAT
 1161
 Ala Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn
 255 260 265 270

CTT CTC CGC CAG GCA CAG GAA AAG TTT GGA AAG GAC AAG TCA CCG AAA
 1209
 Leu Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys
 275 280 285

TTC CAG CTC TTT GGC TCC CCT AGT GGG CAG AAA GAT CTG CTG TTC AAG
 1257
 Phe Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys
 290 295 300

GAC TCT GCC ATT GGG TTT TCG AGG GTG CCC CCG AGG ATA GAT TCT GGG
 1305
 Asp Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly
 305 310 315

CTG TAC CTT GGC TCC GGC TAC TTC ACT GCC ATC CAG AAC TTG AGG AAA
 1353
 Leu Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys
 320 325 330

AGT GAG GAG GAA GTG GCT GCC CGG CGT GCG CGG GTC GTG TGG TGT GCG
 1401
 Ser Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala
 335 340 345 350

GTG GGC GAG CAG GAG CTG CGC AAG TGT AAC CAG TGG AGT GGC TTG AGC
 1449
 Val Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser
 355 360 365

GAA GGC AGC GTG ACC TGC TCC TCG GCC TCC ACC ACA GAG GAC TGC ATC
 1497
 Glu Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile
 370 375 380

GCC CTG GTG CTG AAA GGA GAA GCT GAT GCC ATG AGT TTG GAT GGA GGA
 1545
 Ala Leu Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly
 385 390 395

TAT GTG TAC ACT GCA TGC AAA TGT GGT TTG GTG CCT GTC CTG GCA GAG
 1593
 Tyr Val Tyr Thr Ala Cys Lys Cys Gly Leu Val Pro Val Leu Ala Glu
 400 405 410

AAC TAC AAA TCC CAA CAA AGC AGT GAC CCT GAT CCT AAC TGT GTG GAT
 1641
 Asn Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val Asp

3/28/97

139

415

420

425

430

AGA CCT GTG GAA GGA TAT CTT GCT GTG GCG GTG GTT AGG AGA TCA GAC
1689

Arg Pro Val Glu Gly Tyr Leu Ala Val Ala Val Val Arg Arg Ser Asp
435 440 445

ACT AGC CTT ACC TGG AAC TCT GTG AAA GGC AAG AAG TCC TGC CAC ACC
1737

Thr Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His Thr
450 455 460

3/28/97

GCC GTG GAC AGG ACT GCA GGC TGG AAT ATC CCC ATG GGC CTG CTC TCC
1785

Ala Val Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly Leu Leu Ser
465 470 475

AAC CAG ACG GGC TCC TGC AAA TTT GAT GAA TAT TTC AGT CAA AGC TGT
1833

Asn Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser Cys
480 485 490

GCC CCT GGG TCT GAC CCG AGA TCT AAT CTC TGT GCT CTG TGT ATT GGC
1881

Ala Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile Gly
495 500 505 510

GAC GAG CAG GGT GAG AAT AAG TGC GTG CCC AAC AGC AAC GAG AGA TAC
1929

Asp Glu Gln Gly Glu Asn Lys Cys Val Pro Asn Ser Asn Glu Arg Tyr
515 520 525

TAC GGC TAC ACT GGG GCT TTC CGG TGC CTG GCT GAG AAT GCT GGA GAC
1977

Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asn Ala Gly Asp
530 535 540

GTT GCA TTT GTG AAA GAT GTC ACT GTC TTG CAG AAC ACT GAT GGA AAT
2025

Val Ala Phe Val Lys Asp Val Thr Val Leu Gln Asn Thr Asp Gly Asn
545 550 555

AAC AAT GAG GCA TGG GCT AAG GAT TTG AAC CTG GCA GAC TTT GCG CTG
2073

Asn Asn Glu Ala Trp Ala Lys Asp Leu Asn Leu Ala Asp Phe Ala Leu
560 565 570

CTG TGC CTC GAT GGC AAA CGG AAG CCT GTG ACT GAC GCT AGA AGC TGC
2121

Leu Cys Leu Asp Gly Lys Arg Lys Pro Val Thr Asp Ala Arg Ser Cys
575 580 585 590

CAT CTT GCC ATG GCC CCG AAT CAT GCC GTG GTG TCT CGG ATG GAT AAG
2169

His Leu Ala Met Ala Pro Asn His Ala Val Val Ser Arg Met Asp Lys
595 600 605

GTG GAA CGC CTG AAA CAG GTG CTG CTC CAC CAA CAG GCT AAA TTT GGG
2217

Val Glu Arg Leu Lys Gln Val Leu Leu His Gln Gln Ala Lys Phe Gly
610 615 620

AGA AAT GGA TCT GAC TGC CCG CAG AAG TTT TGC TTA TTC CAG TCT GAA
2265

Arg Asn Gly Ser Asp Cys Pro Gln Lys Phe Cys Leu Phe Gln Ser Glu

3/28/97

141

625	630	635
ACC AAA AAC CTT CTG TTC AAT GAC AAC ACT GAG TGT CTG GCC AGA CTC 2313		
Thr Lys Asn Leu Leu Phe Asn Asp Asn Thr Glu Cys Leu Ala Arg Leu 640	645	650
CAT GGC AAA ACA ACA TAT GAA AAA TAT TTG GGA CCA CAG TAT GTC GCA 2361		
His Gly Lys Thr Thr Tyr Glu Lys Tyr Leu Gly Pro Gln Tyr Val Ala 655	660	665 670

3/28/97

142

GGC ATT ACT AAT CTG AAA AAG TGC TCA ACC TCC CCC CTC CTG GAA GCC
2409

Gly Ile Thr Asn Leu Lys Lys Cys Ser Thr Ser Pro Leu Leu Glu Ala
675 680 685

TGT GAA TTC CTC AGG AAG TAAACCGAA GAAGATGGCC CAGCTCCCCA
2457

Cys Glu Phe Leu Arg Lys
690

AGAAAGCCTC AGCCATTCAC TGCCCCCAGC TCTTCTCCCC AGGTGTGTTG GGGCCTTGGC
2517

TCCCCTGCTG AAGGTGGGGA TTGCCCATCC ATCTGCTTAC AATTCCCTGC TGTCGTCTTA
2577

GCAAGAAGTA AAATGAGAAA TTTTGTTGAT ATTCAAAAAA AA
2619

3/28/97

143

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 711 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Leu	Val	Phe	Leu	Val	Leu	Leu	Phe	Leu	Gly	Ala	Leu	Gly	Leu	-19	-15	-10	-5
Cys	Leu	Ala	Gly	Arg	Arg	Arg	Arg	Ser	Val	Gln	Trp	Cys	Ala	Val	Ser	1	5	10	
Gln	Pro	Glu	Ala	Thr	Lys	Cys	Phe	Gln	Trp	Gln	Arg	Asn	Met	Arg	Lys	15	20	25	
Val	Arg	Gly	Pro	Pro	Val	Ser	Cys	Ile	Lys	Arg	Asp	Ser	Pro	Ile	Gln	30	35	40	45
Cys	Ile	Gln	Ala	Ile	Ala	Glu	Asn	Arg	Ala	Asp	Ala	Val	Thr	Leu	Asp	50	55	60	
Gly	Gly	Phe	Ile	Tyr	Glu	Ala	Gly	Leu	Ala	Pro	Tyr	Lys	Leu	Arg	Pro	65	70	75	
Val	Ala	Ala	Glu	Val	Tyr	Gly	Thr	Glu	Arg	Gln	Pro	Arg	Thr	His	Tyr	80	85	90	
Tyr	Ala	Val	Ala	Val	Val	Lys	Lys	Gly	Gly	Ser	Phe	Gln	Leu	Asn	Glu	95	100	105	
Leu	Gln	Gly	Leu	Lys	Ser	Cys	His	Thr	Gly	Leu	Arg	Arg	Thr	Ala	Gly	110	115	120	125
Trp	Asn	Val	Pro	Thr	Gly	Thr	Leu	Arg	Pro	Phe	Leu	Asn	Trp	Thr	Gly	130	135	140	
Pro	Pro	Glu	Pro	Ile	Glu	Ala	Ala	Val	Ala	Arg	Phe	Phe	Ser	Ala	Ser	145	150	155	
Cys	Val	Pro	Gly	Ala	Asp	Lys	Gly	Gln	Phe	Pro	Asn	Leu	Cys	Arg	Leu	160	165	170	
Cys	Ala	Gly	Thr	Gly	Glu	Asn	Lys	Cys	Ala	Phe	Ser	Ser	Gln	Glu	Pro	175	180	185	
Tyr	Phe	Ser	Tyr	Ser	Gly	Ala	Phe	Lys	Cys	Leu	Arg	Asp	Gly	Ala	Gly	190	195	200	205

3/28/97

1432

Asp	Val	Ala	Phe	Ile	Arg	Glu	Ser	Thr	Val	Phe	Glu	Asp	Leu	Ser	Asp	210	215	220
Glu	Ala	Glu	Arg	Asp	Glu	Tyr	Glu	Leu	Leu	Cys	Pro	Asp	Asn	Thr	Arg	225	230	235
Lys	Pro	Val	Asp	Lys	Phe	Lys	Asp	Cys	His	Leu	Ala	Arg	Val	Pro	Ser	240	245	250
His	Ala	Val	Val	Ala	Arg	Ser	Val	Asn	Gly	Lys	Glu	Asp	Ala	Ile	Trp	255	260	265
Asn	Leu	Leu	Arg	Gln	Ala	Gln	Glu	Lys	Phe	Gly	Lys	Asp	Lys	Ser	Pro	270	275	280
Lys	Phe	Gln	Leu	Phe	Gly	Ser	Pro	Ser	Gly	Gln	Lys	Asp	Leu	Leu	Phe	290	295	300
Lys	Asp	Ser	Ala	Ile	Gly	Phe	Ser	Arg	Val	Pro	Pro	Arg	Ile	Asp	Ser	305	310	315
Gly	Leu	Tyr	Leu	Gly	Ser	Gly	Tyr	Phe	Thr	Ala	Ile	Gln	Asn	Leu	Arg	320	325	330
Lys	Ser	Glu	Glu	Glu	Val	Ala	Ala	Arg	Arg	Ala	Arg	Val	Val	Trp	Cys	335	340	345
Ala	Val	Gly	Glu	Gln	Glu	Leu	Arg	Lys	Cys	Asn	Gln	Trp	Ser	Gly	Leu	350	355	360
Ser	Glu	Gly	Ser	Val	Thr	Cys	Ser	Ser	Ala	Ser	Thr	Thr	Glu	Asp	Cys	370	375	380
Ile	Ala	Leu	Val	Leu	Lys	Gly	Glu	Ala	Asp	Ala	Met	Ser	Leu	Asp	Gly	385	390	395
Gly	Tyr	Val	Tyr	Thr	Ala	Cys	Lys	Cys	Gly	Leu	Val	Pro	Val	Leu	Ala	400	405	410
Glu	Asn	Tyr	Lys	Ser	Gln	Gln	Ser	Ser	Asp	Pro	Asp	Pro	Asn	Cys	Val	415	420	425
Asp	Arg	Pro	Val	Glu	Gly	Tyr	Leu	Ala	Val	Ala	Val	Val	Arg	Arg	Ser	430	435	440
Asp	Thr	Ser	Leu	Thr	Trp	Asn	Ser	Val	Lys	Gly	Lys	Lys	Ser	Cys	His	450	455	460
Thr	Ala	Val	Asp	Arg	Thr	Ala	Gly	Trp	Asn	Ile	Pro	Met	Gly	Leu	Leu	465	470	475
Ser	Asn	Gln	Thr	Gly	Ser	Cys	Lys	Phe	Asp	Glu	Tyr	Phe	Ser	Gln	Ser	480	485	490

146

3/28/97

144

Cys	Ala	Pro	Gly	Ser	Asp	Pro	Arg	Ser	Asn	Leu	Cys	Ala	Leu	Cys	Ile
495						500					505				
Gly	Asp	Glu	Gln	Gly	Glu	Asn	Lys	Cys	Val	Pro	Asn	Ser	Asn	Glu	Arg
510					515					520					525
Tyr	Tyr	Gly	Tyr	Thr	Gly	Ala	Phe	Arg	Cys	Leu	Ala	Glu	Asn	Ala	Gly
				530					535					540	
Asp	Val	Ala	Phe	Val	Lys	Asp	Val	Thr	Val	Leu	Gln	Asn	Thr	Asp	Gly
			545					550					555		
Asn	Asn	Asn	Glu	Ala	Trp	Ala	Lys	Asp	Leu	Asn	Leu	Ala	Asp	Phe	Ala
		560					565					570			
Leu	Leu	Cys	Leu	Asp	Gly	Lys	Arg	Lys	Pro	Val	Thr	Asp	Ala	Arg	Ser
	575					580					585				
Cys	His	Leu	Ala	Met	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Met	Asp
590					595					600					605
Lys	Val	Glu	Arg	Leu	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Lys	Phe
				610					615					620	
Gly	Arg	Asn	Gly	Ser	Asp	Cys	Pro	Gln	Lys	Phe	Cys	Leu	Phe	Gln	Ser
			625					630					635		
Glu	Thr	Lys	Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Arg
		640					645					650			
Leu	His	Gly	Lys	Thr	Thr	Tyr	Glu	Lys	Tyr	Leu	Gly	Pro	Gln	Tyr	Val
	655					660					665				
Ala	Gly	Ile	Thr	Asn	Leu	Lys	Lys	Cys	Ser	Thr	Ser	Pro	Leu	Leu	Glu
670					675					680					685
Ala	Cys	Glu	Phe	Leu	Arg	Lys									
				690											

3/28/97

145

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCCATGGGGG TCACAAAGAA CTGGAC

26

3/28/97

146

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGAAGCTTGC TAACAGTATA TCATAGG

27

3/28/97

147

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGGGACTCC ACAGTTATGG

20

3/28/97 148

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCACACAATT ATTTGATATG

20

3/28/97

149

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTTGCTGTGG CGGTGGTTAG GAGATCAGAC

30

3/28/97

150

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTCCTGGAAG CCTGTGAATT CTCAGGAAG

30

3/28/97

151

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACCAAGTGCT TCCAGTGGCA G

21

154

3/28/97 152

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCCATGGGGG TCACAAAGAA CTGGAC

26

3/28/97

153

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGAAGCTTGC TAACAGTATA TCATAGG

27

3/28/97

154

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GAGGGACTCC ACAGTTATGG

20

3/28/97

155

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCACACAATT ATTTGATATG

20

158

3/28/97

156

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATGAAACTTA TCCTCACCTG TCTTGTG

27

159

3/28/97

157

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGGTTTTCGA GGGTGCCCC GAGGATGGAT

30

160

3/28/97 158

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGTCGACAGT AC

12

3/28/97

159

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGTCGACGGT AC

12

162

3/28/97 160

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGACGTTGTA AAACGACGG

19

3/28/97

141

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATTGTCGACT TATCGATGGG TTGATGATCA AGGTGA

36

164

3/28/97

162

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CAAATCGATT GAACTTGCAG TATCTCCACG AC

32

165

3/28/97

163

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGGATCGATC AGATTCTGTC CCCCAT

26

1166

3/28/97

164

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGATCCGAGA CACAGAACAG G

21

167

3/28/97 165

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCTAATCCAT CCATCCTATA G

21

3/28/97

166

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TTTGGAAGG ACAAGTCACC G

21

3/28/97

167

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTCACTTTTC CTCAAGTTCT G

21

3/28/97

168

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 807 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGAAGTGCCT GGAGATTAAA ATGTGAGAGT GGAGTGGAGG TTGGGTCCTG TAGGCCTTCC
60
CATCCACGT GCCTCACGGA GCCCTAGTGC TACTCAGTCA TGCCCCGCA GCAGGGGTCA
120
GGTCACTTTC CCATCCTGGG GGTTATTATG ACTGTTGTCA TTGTTGTTGC CATTTTGTCT
180
ACCCTAACTG GGCAGCGGGT GCTTGCAGAG CCCTCGATAC TGACCAGGTT CCCCCTCGG
240
AGCTCGACCT GAACCCCATG TCACCCTCGC CCCAGCCTGC AGAGGGTGGG TGACTGCAGA
300
GATCCCTTTA CCCAAGGCCA CAGTCACATG GTTTGGAGGA GATGGTGCCC AAGGCAGAAG
360
CCACCCTCCA GACACACCTG CCCCAGTGC TGGCTCTGAC CTGTCCTTGT CTAAGAGGCT
420
GACCCAGAA GTGTTCTTGG CGCTGGCAGC CAGCCTGGAC CCAGAGCCTG GACACCCCT
480
GCGCCCCAC TTCTGGGGG GTACCAGGAA CCGTCCAGGC CCAGAGGGCC TTCCTGCTTG
540
GCCTCGAATG GAAGAAGGCC TCCTATTGTC CTTCGTAGAG GAAGCAACCC CAGGGCCCAA
600
GGATAGGCCA GGGGGGATTC GGGGAACCGC GTGGCTCCGG CGCGGCCCGG GCTGGCTGGC
660
TGGCCCTCCT CCTGTATAAG GCCCCGAGCC CGCTGTCTCA GCCCTCCACT CCCTGCAGAG
720
CTCAGAAGCG TGACCCAGC TGCAGCCATG AAGTGCCTCC TGCTTGCCCT GGCCCTCACC
780

3/28/97

1682

TGTGGCGCCC AGGCCCTCAT CGTCACC
807

3/28/97

169

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 824 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGAAGTGTCC TGGGAGATTT AAAATGTGAG AGGCGGGAGG TGGGAGGTTG GGCCCTGTGG
60
GCCTGCCCAT CCCACGTGCC TGCATTAGCC CCAGTGCTGC TCAGCCGTGC CCCC GCCGCA
120
GGGGTCAGGT CACTTCCCCG TCCTGGGGTT ATTATGACTC TTGTCATTGC CATTGCCATT
180
TTTGCTACCC TAACTGGGCA GCAGGTGCTT GCAGAGCCCT CGATACCGAC CAGGTCCTCC
240
CTCGGAGCTC GACCTGAACC CCATGTCACC CTTGCCCCAG CCTGCAGAGG GTGGGTGACT
300
GCAGAGATCC CTTACCCAA GGCCACGGTC ACATGGTTTG GAGGAGCTGG TGCCCAAGGC
360
AGAGGCCACC CTCCAGGACA CACCTGTCCC CAGTGCTGGC TCTGACCTGT CTTGTCTAA
420
GAGGCTGACC CCGGAAGTGT TCCTGGCACT GGCAGCCAGC CTGGACCCAG AGTCCAGACA
480
CCCACCTGTG CCCCCGCTTC TGGGGTCTAC CAGGAACCGT CTAGGCCAG AGGGGGACTT
540
CCTGCTTGGC CTTGGATGGA AGAAGGCCTC CTATTGTCCT CGTAGAGGAA GCCACCCCGG
600
GGCCTGAGGA TGAGCCAAGT GGGATTCCGG GAACCGCGTG GCTGGGGGCC CAGCCCGGGC
660
TGGCTGGCCT GCATGCGCCT CCTGTATAAG GCCCAAGCC TGCCTGTCTC AGCCCTCCAC
720
TCCCTGCAGA GCTCAGAAGC ACGACCCAG CTGCAGCCAT GAAGTGCCTC CTGCTTGCCC
780

3/28/97

170

TGGGCCTGGC CCTCGCCTGT GCGTCCAGG CCATCATCGT CACC
824

3/28/97

171

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATCACCTTGA TCATCAACCC AGCTTGCTGC TTCTTCCCAG TCTTGGGTTC AAGGTATTAT
60

GTATACATAT AACAAAATTT CTATGATTTT CCTATGTCTC ATCTTTCATT CTTCACTAAT
120

ACGCAGTTGT AACTTTTCTA TGTGATTGCA AGTATTGGTA CTTTCCTATG ATATACTGTT
180

AGCAAGCTTG AGGTGTGGCA GGCTTGAGAT CTGGCCATAC ACTTGAGTGA CAATGACATC
240

CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCAG
288

3/28/97 172

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..3, 64..68
- (D) OTHER INFORMATION: /note= "Overhang"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGATACCAAG TCGCCTCCAG ACCGCAGACA TGAAACTTGT CTCCTCGTC CTGCTGTTCC
60

TCGGGGGCC
68

3/28/97

173

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GACTGTGTCT GGCT

14

3/28/97

174

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..2
- (D) OTHER INFORMATION: /note= "Overhang"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 17..20
- (D) OTHER INFORMATION: /note= "Overhang on complementary

strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CGGTCGACAT CGATGC

16

3/28/97

175

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: -1..-4, 31..34
- (D) OTHER INFORMATION: /note= "Overhangs on complementary

strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAGAAGCGTG ACCCCAGTAT CGATACCTGG

30

3/28/97

176

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: -1..-4,
- (D) OTHER INFORMATION: /note= "Overhang on complementary strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CATCGATCCC TAGCACTCTG ACCTAGCAGT C

31

180

3/28/97

1.77

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..4
- (D) OTHER INFORMATION: /note= "Overhang"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 39..42
- (D) OTHER INFORMATION: /note= "Overhang on complementary

strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TCGAGCGGCC GCCGGACCGG GCCGCCTCGG CCTCGCGA

38

3/28/97

178

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..2
- (D) OTHER INFORMATION: /note= "Overhang"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CGATAACCAT GAAACTTCTT ATCCTCACCT GTCTTGTGGC TGTTGCTCTT G
51

182

3/28/97

179

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CCAAGGTCTT TGAAAGGTGT GAGTTGC

27